Process Mining: a framework proposal for Pervasive Business Intelligence

Teresa Guarda
Isla - Superior Institute of Languages and Administration
Leiria, Portugal

Manuel Filipe Santos
Minho University
Guimarães, Portugal

Maria Fernanda Augusto
Isla - Superior Institute of Languages and Administration
Leiria, Portugal

Carlos Silva
Isla - Superior Institute of Languages and Administration
and CIEO - Centre for Spatial and Organizational Dynamics
Leiria, Portugal

Filipe Pinto
Polytechnic Institute of Leiria
Leiria, Portugal

Abstract— **In recent years, global growth slowed, the markets have matured and become more competitive. The impact of computing in organizations made information technology a strategic element to the acquisition and maintenance of competitive advantage. Based on the literature review in the related areas of Business Intelligence (BI) and Process Mining (PM), is presented a framework for improving the decision-making processes in organizations.**

Keywords- Business Intelligence; Pervsive Business Intelligence ; Process Mining; Competitive Advantage

I. INTRODUCTION

Business Intelligence promotes an environment that facilitates access to information needed for daily activities o the organization, allowing analyzing the current situation of the business and its performance. The systems and BI tools play a key role in the strategic planning process of organizations. Information processing has gradually become the basis for achieving competitive advantage. The organization has to believe it has the right information at the right time and available to the right people [1].

Initially one of the purposes of the BI was the easy distribution of information to the no specialist end user. For implement in small and medium enterprises (SME) there are increased difficulties, BI is not easy to implement in SME due to the following factors [2]: high prices; demanding requirements in terms of hardware and software infrastructure; complexity use for most users; irrelevant functionalities; low flexibility to deal with a wild change business environment; and not given due attention to the different needs of data access of SMEs and large firms.

Nowadays, the number of organizations that provide BI to their decision makers is increasing. It is increasingly important and difficult for organizations make the right decisions for the business. Companies know that the ability to make the right decisions frequently is essential for increased profits; for risk management and for good overall performance. Due to uncontrollable factors, such as fast-moving markets, the economic and the regulation changes, and new sources of competition, making the right decision is not peaceful.

This paper is organized as follows: after this introductory part we present related background concepts. Then, the main contribution is presented in terms of a framework proposal for PM integration on Pervasive Business Intelligence (PBI). Finally we draw some conclusions.

II. BACKGROUND

A. Business Intelligence

The exponential growth of data volumes generated by organizations forced the development of BI applications very sophisticated, and also expanded the type of tools that classic BI systems normally use. The BI market is growing, and large companies still need new solutions to new business needs.

Stackowiak et al. was defined BI as the process of collecting large amounts of data, its analysis, and consequent production of reports that summarize the essence of the business data actions that will assist the manager in business decision-making [3]. For Zeng et al. BI is the process of collection, processing and dissemination of information, in order to reduce uncertainty in strategic decision making [4]. The BI is the process by which users obtain accurate and consistent data from the storage environment of organizational data (data warehouses, data marts). Data obtained from various business contexts, allow users to identify, analyze and detect trends, anomalies, and make predictions. The systems and BI tools play a key role in the strategic planning process of

organizations. These systems allow collect, store, access and analyze organizational data in order to support and facilitate the decision making [3].

The BI has several approaches, among these we may refer: the management approach; the technique approach and the value added approach. The management approach views the BI as the process in which data collected within and outside the organization are integrated in order to create relevant knowledge to the decision-making process, emphasizing the management decision-making [4, 5, 6, 7, 8, 9]. On the other hand the technical approach presents the BI as a set of tools that support the process shown in management approach, with the emphasis on the technology used, and not in the process [10, 11, 12, 13, 14]. While for the value-added approach, BI offers added value to achieve competitive advantage [15, 16, 5, 17, 18].

BI uses several tools, like Data Warehouses (DM), Data Mining (DM), OLAP tools, among others, that will be used in the collection, analysis and extraction of information that will help the process of management and decision making.

Gartner recognizes five levels of maturity for BI: unconscious, tactical, focused, strategic, and pervasive [21]. The main features of the maturity levels are [22]:

- Unconscious: Level characterized by inconsistent data, data interpretation incorrect and inconsistent, and constant changes either for the needs of individual or departmental information. The use of spreadsheet is high and the use of communication tools is limited. The information management and communication are left to the IT department.
- Tactical: In this level organizations begin to invest in BI. Choose software that does not respond to their needs. Most users are not sufficiently qualified to take advantage of the system. The management does not trust in the quality and consistency of information provided.
- Focused: Is the level at which the organization has the first success, and that BI brings some benefits to the business. The metrics are used mostly only at the departments level. The aim is to optimize the efficiency of the departments or business units, with no relationship to the overall business objectives. The inconsistencies of metrics are common. The data are not integrated at this stage and are available through applications covering only part of the business.
- Strategic: Organizations now have a clear strategy for the development of BI. Some organizations choose to include the BI and business process management in critical business processes. The information is available to all employees of the organization. The policy of data management and data quality metrics are implemented. Data quality is under constant supervision. The strategic information becomes reliable and is used as support for strategic decision making. Users are trained in data processing and know how to use them effectively in strategic and tactical decisions.
- Pervasive: BI is pervasive across all areas of business and cross referenced with the organizational culture. BI systems become part of business processes, providing greater flexibility in the adaptation to the changes of business needs and information.

The higher the maturity level of BI, greater the value to the business.

B. Pervasive Business Intelligence

In the current scenario of economic recession, the departments of information technology (IT) are being reevaluated, being forced to become more efficient [23], what has led organizations looking for alternative ways to increase the value of their BI initiatives. In this sense, it is increasing the effort of organizations to achieve pervasive BI [24].

The Pervasive Business Intelligence (PBI) appears as the natural evolution of Business Intelligence applications in organizations, presenting two directions, vertical (top-down) and horizontal (across departments), with application from strategic to operational level.

The PBI appears on the last level of BI maturity. There are several definitions of PBI, is the ability to deliver timely manner to all users, the integrated information in data warehouses, providing the necessary visibility, knowledge, and facts for decision making in all business processes [25]; is the improvement of the capabilities of strategic and operational decision-making of an organization, through the design and implementation of organizational culture, business processes and technology as a whole [26], is the BI for the whole organization, available to all people, at all levels of the organization, analysis, alerts and feedback mechanisms [27].

PBI implementation is supported by applications that allow access the data in real time, supporting the actions of CRM and marketing campaigns. The application of PBI is reinforced when front line employees are on contact with customers and can create new sales opportunities, up-sell and cross-sell [25]. The PBI aims to align all processes to allow the delivery of relevant information to users who need support in decision making.

The PBI is the operationalization of BI across the organization, enabling BI systems reaches all levels of the organization, at the right time and with the information needed.

C. Process Mining

To maintain and build competitive advantage, organizations need to adapt constantly changing business processes to meet the needs and expectations of the business, and the changes in the business environment. To improve business processes, enterprises have to make the necessary changes, being necessary redesigning processes. In this context, managers need a method that allows them to determine in a concise and timely manner, the processes that no longer meet the needs of the business and need to be redesigned.

The term process mining refers to the techniques and tools that enable the extraction of knowledge from event logs

available in enterprise information systems. These techniques and tools provide new ways to discover, monitor and improve processes. The PM assumes that it is possible to get the flow of activities for a process from execution logs of transactions undertaken in information systems. The initial process mining techniques allowed reaching satisfactory results for well-structured processes, but failed in the case of ill-structured processes or poorly defined, where there wasn't a strong dependency between activities [19]. Many of the information systems that support business processes record the events associated with the execution of the actual flow of processes (logs). On many of the logs, it is possible to identify patterns that can be used in the discovery of knowledge. Through the logs is possible to identify how the processes are structured, their relationships, and the people involved. The aim of process mining is the extraction of information from the logs of events in order to capture the business process as it was implemented. The process mining has emerged as a way to analyze systems and their effective use, based on log events generated by these systems. The PM has been investigated in recent years [20, 21, 22]. The subject has a close relationship with the classical concept of data mining. According Han et al. [23], data mining refers to the knowledge extraction of large quantities of data, for automatic or semiautomatic means in order to discover meaningful patterns and rules. Similarly, the PM focuses on pattern extraction. However, it is facing the relationship of sequence and competition between events. While data mining focuses on highlighting trends in data and relationships between attributes, PM seeks to understand how a given process is executed, based on the analysis of each one of its occurrences.

In the beginning of the century, there was a considerable increase of scientific development in this area, with several publications, which had contribute to new approaches to PM techniques and tools for its implementation. PM emerges intending to obtain objective information about how the process is run, since it uses actual data from the execution of the process, produces direct and tangible results [24].

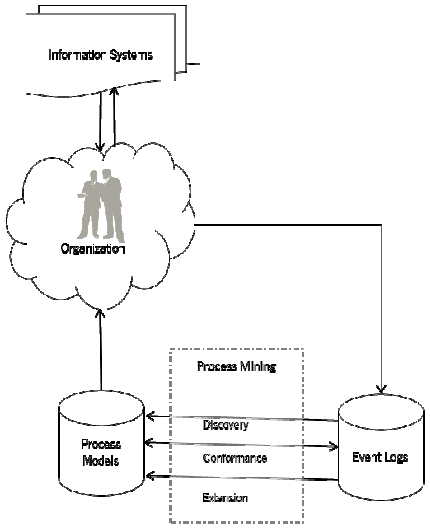

Figure 1. Process Mining (Adapted from van der Aalst [28])

The more complex the process, there will be greater difficulty in understanding them. So when the cases involve distributed activities, is very difficult to the organizations employees, share or have a common vision of processes, especially if they are constantly changing [29].

The PM may be performed in three perspectives: discovery, conformity analysis and extension (Fig.1). The discovery occurs when there is a formal model of the process, and it may be extracted from the data produced during the execution of the process. The conformity analysis presupposes the existence of a previous model, is verified if the process coincides with the model. The compliance analysis allows detection of deviations, and perceives their causes and possible impacts. Finally the extension is the improvement of an existing model based on data analysis of their execution.

III. FRAMEWORK PROPOSAL

The aim of this paper is to integrate the process mining in BI systems through the development of a framework with the goal of assisting in the collection and pre-processing of information and its dissemination, to assist managers in the decision-making process. The construction of the framework is developed through the analysis of key definitions of BI, PM and related concepts. It will be understood as a reactive and dynamic analysis made by PM over running processes, allowing respond to information needs, driven by the restructuring process models, which will enable the data-intensive analysis in the context of BI, with statistical and quantitative analysis, and predictive and explanatory models that assist the manager in decision-making.

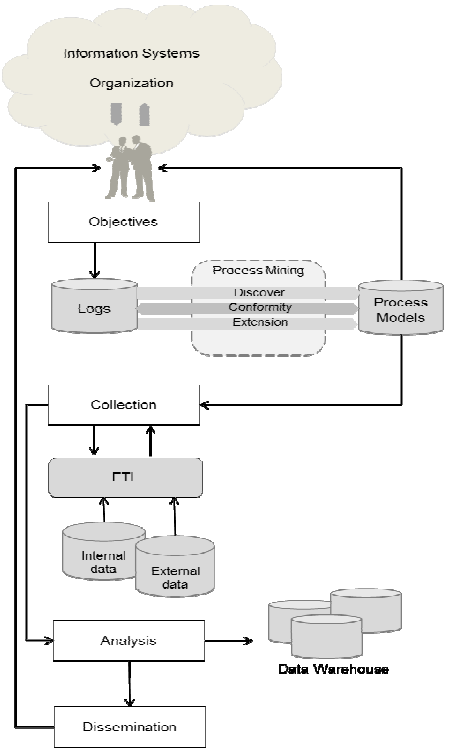

Figure 2. Process Mining framework for PBI

The process for satisfying user's information needs is complex and composed by different activities to be exploited. The propose framework (Fig.3) comprises four basic processes: objectives, collection, analysis and dissemination. Initially are defined the objectives regarding users' information needs for decision-making process. This process is leveraged by the PM for evaluation and suitability of process models to the requirements defined in the initial process. Then the collection process, oriented by the appropriate process model, extract, transforms and load (ETL) organization internal and external data sources. The ETL is a process for extracting data from a database system, where the data are processed, modified, and then inserted into another database. In the next process, analysis, all activities should be developed in order to analyze the data, looking for patterns, and loaded organized and coded information on the data warehouse data marts. In the disseminations process, results will be distributed to end-users for posterior feedback if needed.

IV. CONCLUSIONS

BI can help companies in reducing costs and in the increase revenue through optimizing business processes. PBI empowers peoples at all organization levels, with analytics, alerts and feedback instruments timely. PBI system has to reflect the real time concept to a particular business, i.e., the right time. The right time concept, defines the information to be collected. The PBI framework is a complex process that goes from user's objectives (information needs), collection of internal\external data from the organizations environment, data analysis, until the dissemination of generated quality information to assist decision-making. The overall vision of the organization is achieved through the redesign process models. And this is the key to acquiring or maintaining competitive advantage, allowing the organization to be proactive and reactive.

REFERENCES

[1] A. Palmer e B. Hartley, The Business and Marketing Environment, 3rd ed., London: McGraw-Hill, 2000.

[2] G. Xie, Y. Yang, S. Liu, Z. Qiu, Y. Pan e X. Zhou, "EIAW: towards a business-friendly data warehouse using semantic web technologies," The Semantic Web, pp. 857-870, 2007.

[3] R. R. J. Stackowiak e R. Greenwald, Oracle Data Warehousing and Business Intelligence Solutions, Indianapolis: Wiley Publishing, 2007.

[4] L. Zeng, L. Xu, Z. Shi, M. Wang e W. Wu, "Systems, Man and Cybernetics, 2006. SMC'06. IEEE International Conference," em IEEE International Conference, 2006.

[5] D. A. Aaker, V. Kumar, G. S. Day e R. Leone, Marketing Research, 10th Edition, Wiley.com, 2009.

[6] M. Petrini e M. Pozzebon, "What role is Business Intelligence playing in developing countries. A picture of Brazilian companies," Cahier du GReSI no, pp. 2-16, 2004.

[7] D. J. Power, "Understanding Data-Driven Decision Support Systems," Information Systems Management, pp. 149-154, 2008.

[8] T. Bucher, A. Gericke e S. Sigg, "Process-centric business intelligence," Business Process Management Journal, pp. 408-429, 2009.

[9] H. Cheng, Y. Lu e C. Sheu, "An ontology-based business intelligence application in a financial knowledge management system," Expert Systems with Applications, pp. 3614-3622, 2009.

[10] R. Kalahasthi e P. Krishna, "Cell Based Reporting For Effective Custom Reports And Analytics," Business Intelligence Journal, pp. 425-430, 2009.

[11] A. H. Lim e C.-S. Lee, "Processing online analytics with classification and association rule mining," Knowledge-Based Systems, pp. 248-255, 2010.

[12] T. Back, "Adaptive business intelligence based on evolution strategies: some application examples of self-adaptive software," Information Sciences, pp. 1-9, 2002.

[13] L. Van Dyk e P. Conradie, "Creating business intelligence from course management systems," Campus-Wide Information Systems, pp. 120-133, 2007.

[14] H. Baars e H. Kemper, "Management Support with Structured and Unstructured Data: An Integrated Business Intelligence Framework," Information Systems Management, pp. 132-148, 2008.

[15] B. Sahay e J. Ranjan, "Real time business intelligence in supply chain analytics," Information Management & Computer Security, pp. 28-48, 2008.

[16] M. Chen e S. Wang, "The use of a hybrid fuzzy-Delphi-AHP approach to develop global business intelligence for information service firms," Expert Systems with Applications, pp. 7394-7407, 2010.

[17] H. Wang e S. Wang, "A knowledge management approach to data mining process for business intelligence," Industrial Management & Data Systems, pp. 622-634, 2008.

[18] P. Trim e Y. Lee, "A strategic marketing intelligence and multi-organisational resilience framework," European Journal of Marketing, pp. 731-745, 2008.

[19] C. Fleisher, "Using open source data in developing competitive and marketing intelligence," European Journal of Marketing, pp. 852-866, 2008.

[20] T. Davenport, J. Harris e J. Shapiro, "Competing on talent analytics," Harvard Business Review, pp. 52-58, 2010.

[21] B. Burtin, "Toolkit: Maturity Checklist for Business Intelligence and," Gartner Inc. Research, 2007.

[22] N. Rayner e K. Schlegel, "Maturity Model Overview for Business Intelligence and Performance Management," Gartner Inc. Research, 2008.

[23] C. Imhoff e C. White, "Self-Service Business Intelligence: Empowering Users To Generate Insights," TDWI Best Practices Report Third Quarter 2011, 2011.

[24] S. Ortiz, "Taking Business Intelligence to the Masses," Computer, vol. 43, pp. 12-15, 2010.

[25] J. Markarian, S. Brobst e J. Bedell, "Critical Success Factors Deploying Pervasive BI," Informatica, Teradata, MicroStrategy, 2008.

[26] D. Vesset e B. McDonough, Improving Organizational Performance Management Through, IDC, 2009.

[27] D. Mittlender, "Pervasive Business Intelligence: Enhancing Key Performance Indicators," DM Review, 2005.

[28] W. van der Aalst e C. Giinther, "Finding structure in unstructured processes: The case for process mining," em Application of Concurrency to System Design, 2007. ACSD 2007. Seventh International Conference on, Washington, 2007.

[29] W. van der Aalst, B. van Dongen, J. Herbst, L. Maruster, G. Schimm e A. Weijters, "Workflow mining: a survey of issues and approaches," Data and Knowledge Engineering, pp. 237-267, 2003.